01-18-06

AF/1631

Express Mailing Label No.: EV 589980937 US
IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

In re Application of
Douglas M. BLAIR
Appln. No.: 09/881,234
Filing Date: Jun. 14, 2001

Atty. Docket No.: 2551-026

Group Art Unit: 1631
Examiner: Smith, C.

For: **APPARATUS AND METHOD FOR PROVIDING SEQUENCE DATABASE COMPARISON**

* * * * * * * * * * * * * *

TRANSMITTAL LETTER

* * * * * * * * * * * * * *

Mail Stop Appeal Brief - Patents
Commissioner for Patents
P.O. Box 1450
Alexandria, VA 22313-1450

Dear Sir:

Enclosed, please find:

1. Appellants' Amended Brief on Appeal Under 37 C.F.R. §41.47.

The Commissioner is hereby authorized to charge any fee deficiency, or credit any overpayment, to Deposit Account No. 18-1579. However, it is believed that no fees are due with this filing, since Appellant in Application Ser. No. 09/881,234 previously paid the fees under 37 C.F.R. 1.17(b) for a Notice of Appeal on December 9, 2003 and paid the fees under 37 C.F.R. 1.17(c) for a Brief on Appeal on March 9, 2004. The provisions for filing an amended Appeal Brief under 37 C.F.R. 41.37(d) in response to a Notification of Non-Compliant Appeal Brief have no requirement for additional fees.

Respectfully submitted,

Christopher B. Kilner, Reg. No. 45,381
Roberts Abokhair & Mardula, LLC
11800 Sunrise Valley Dr., Suite 1000
Reston, VA 20191
(703) 391-2900

IN THE UNITED STATES PATENT AND TRADEMARK OFFICE
BEFORE THE BOARD OF PATENT APPEALS AND INTERFERENCES

In re Application of Atty. Docket No.: 2551-026

Douglas M. BLAIR

Appln. No.: 09/881,234 Group Art Unit: 1631

Filing Date: Jun. 14, 2001 Examiner: Smith, C.

For: **APPARATUS AND METHOD FOR PROVIDING SEQUENCE DATABASE COMPARISON**

* *

APPELLANTS' AMENDED BRIEF ON APPEAL UNDER 37 C.F.R. § 41.47

* *

Mail Stop Appeal Brief - Patents
Commissioner for Patents
P.O. Box 1450
Alexandria, VA 22313-1450

Dear Sir:

In accordance with the provisions of 37 C.F.R. § 41.37 and the Notification of Non-Complaint Appeal Brief mailed January 3, 2006, Appellant submits the following:

I. REAL PARTY IN INTEREST

Based on information supplied by Appellants, and to the best of Appellants' legal representatives' knowledge, the real party in interest is Parabon Computation, Inc.

II. RELATED APPEALS AND INTERFERENCES

Appellants, as well as Appellants' assigns and legal representatives are unaware of any appeals or interferences which will be directly affected by, or which will directly

affect, or have a bearing on the Board's decision in the pending appeal.

III. STATUS OF CLAIMS

Claims 1-23 are currently pending. No claims have been allowed. No claims have been canceled. Claims 1-23 are appealed. Claims 1-23, as amended with the original Appeal Brief of March 9, 2004, are set forth in the Appendix.

IV. STATUS OF AMENDMENTS

An amendment was filed with the original Appeal Brief of March 9, 2004 to eliminate the alleged indefiniteness of the abbreviations CPU, ID, and BLAST from the claims so as to place the claims in better condition for appeal and was entered.

V. SUMMARY OF CLAIMED SUBJECT MATTER

Appellants' claimed invention is directed to a method and system of comparing a query and a subject database using a distributed computing platform. The databases are divided into data elements having a size within a specified range. All data elements and task definitions are sent to a master CPU of a master-slave distributed computing platform, wherein task definitions comprise at least one comparison parameter, at least one executable comparison element, and a query and a subject data element ID/descriptor. Data elements are sent alternately from query and subject data elements. A task definition is sent for each task from the master CPU to one of a plurality of slave CPUs when all parts of a task definition and data

elements referenced by the task definition are available at the master CPU. Data elements are then sent to the slave CPUs for performance of the tasks. Task results for each task are returned to a CPU.

In the embodiment of independent claim 1, the invention is drawn to a method of comparing a query dataset N with a subject dataset M, comprising: dividing said query dataset N into n_N data elements having a size within a specified range and dividing said subject dataset M into n_M data elements having a size within said specified range (see page 18, line 9 to page 19, line 15, page 28, line 18 to page 29, line 3, figs. 1B, 3, and box 620 of figure 6); determining a number of tasks for an entire comparison of datasets N and M as $n_N \times n_M$ (see page 19, lines 16-20, page 29, line 19 to page 30, line 1, box 628 of fig. 6); sending all data elements and task definitions to a master CPU of a master-slave distributed computing platform, wherein task definitions comprise at least one comparison parameter, at least one executable element capable of performing comparisons, a query data element ID/descriptor, and a subject data element ID/descriptor, and wherein data elements are sent alternately from query and subject data elements (see page 19, line 21 to page 20, line 14, page 29, line 19 to page 30, line 19, computer topology of fig. 5 and boxes 630-640 of fig. 6); sending a task definition for each task from the master CPU to one of a plurality of slave CPUs when all parts of a task definition and data elements referenced by said task definition are available at said master CPU and sending data elements referenced by said task definition to said slave CPU (see page 21, lines 15-18, page 30, lines 9-19, box 650 of fig. 6); performing

each task on a slave CPU (see page 21, lines 15-18, page 30, lines 9-19, box 650 of fig. 6); and returning task results for each task to said master CPU (see page 21, line 19 to page 23, line 7, page 32, lines 1-3, box 730 of fig. 7).

In the embodiment of independent claim 13, the invention is drawn to a system for comparing a query dataset N with a subject dataset M, comprising: a master CPU of a master-slave distributed computing platform; a plurality of slave CPUs capable of communication with said master CPU; and a client CPU (see page 17, lines 9-20, fig. 5) with instructions for: dividing said query dataset N into n_N data elements having a size within a specified range and dividing said subject dataset M into n_M data elements having a size within said specified range (see page 18, line 9 to page 19, line 15, page 28, line 18 to page 29, line 3, figs. 1B, 3, and box 620 of figure 6); determining a number of tasks for an entire comparison of datasets N and M as $n_N \times n_M$ (see page 19, lines 16-20, page 29, line 19 to page 30, line 1, box 628 of fig. 6); sending all data elements and task definitions to said master CPU of a master-slave distributed computing platform, wherein task definitions comprise at least one comparison parameter, at least one executable element capable of performing comparisons, a query data element ID/descriptor, and a subject data element ID/descriptor, and wherein data elements are sent alternately from query and subject data elements (see page 19, line 21 to page 20, line 14, page 29, line 19 to page 30, line 19, computer topology of fig. 5 and boxes 630-640 of fig. 6); said master CPU comprising instructions for: sending a task definition for each task

to one of said plurality of slave CPUs when all parts of a task definition and data elements referenced by said task definition are available at said master CPU; and sending data elements referenced by said task definition to said slave CPU (see page 21, lines 15-18, page 30, lines 9-19, box 650 of fig. 6); and said slave CPUs including instructions for: performing each task (see page 21, lines 15-18, page 30, lines 9-19, box 650 of fig. 6); and returning task results for each task to said master CPU (see page 21, line 19 to page 23, line 7, page 30, line 21 to page 31, line 23, box 730 of fig. 7).

VI. GROUNDS OF REJECTION TO BE REVIEWED ON APPEAL

The rejections of claims 4, 7, 8, 16, and 19 in the Final Office Action of January 21, 2005, as indefinite under the second paragraph of 35 U.S.C. § 112, were withdrawn by the Examiner at page 11 of the Examiner's Answer mailed July 13, 2005, and therefore are no longer grounds of rejection to be reviewed on appeal.

The grounds of rejection to be reviewed on Appeal are:

Grounds 1 - Are claims 1 and 13, and all the remaining claims dependent thereon, indefinite under the second paragraph of 35 U.S.C. § 112 due to the term "said task definition" lacking a proper antecedent basis in the claims?

Grounds 2 - Are claims 1, 4, 6-13, and 18-23 unpatentable over the publication to Smith et al. (1996) in view of the publication to Altschul et al. (1990) and U.S. Patent No. 5,862,325 to Reed et al. as being obvious?

VII. ARGUMENT

Grounds 1

Claim Rejections - 35 USC §112

As previously submitted and cited in M.P.E.P. §2173.01, Appellants submit that a fundamental principle contained in the second paragraph of 35 U.S.C. § 112 is that Appellants are their own lexicographers. Appellants can define in the claims what they regard as their invention essentially in whatever terms they choose so long as the terms are not used in ways that are contrary to accepted meanings in the art. Appellants may use functional language, alternative expressions, negative limitations, or any style of expression or format of claim which makes clear the boundaries of the subject matter for which protection is sought. As noted by the court in *In re Swinehart*, 439 F.2d 210, 160 USPQ 226 (CCPA 1971), a claim may not be rejected solely because of the type of language used to define the subject matter for which patent protection is sought.

Appellants again submit that the proper focus during examination of claims for compliance with the requirement for definiteness of 35 U.S.C. §112, second paragraph as defined in M.P.E.P. §2173.02 is whether the claim meets the threshold requirements of clarity and precision, not whether more suitable language or modes of expression are available. When the Examiner is satisfied that patentable subject matter is disclosed, and it is apparent to the examiner that the claims are directed to such patentable subject matter, he or she should allow claims which define the patentable subject matter with a reasonable degree of particularity and distinctness. *Some latitude in the manner*

of expression and the aptness of terms should be permitted even though the claim language is not as precise as the examiner might desire. Examiners are encouraged to suggest claim language to appellants to improve the clarity or precision of the language used, *but should not reject claims or insist on their own preferences if other modes of expression selected by appellants satisfy the statutory requirement.*

The essential inquiry pertaining to this requirement is whether the claims set out and circumscribe a particular subject matter with a reasonable degree of clarity and particularity. Definiteness of claim language must be analyzed, not in a vacuum, but in light of:

(A) The content of the particular application disclosure;

(B) The teachings of the prior art; and

(C) The claim interpretation that would be given by one possessing the ordinary level of skill in the pertinent art at the time the invention was made.

Claims 1 and 13

Claims 1 and 13 were rejected as indefinite under the second paragraph of 35 U.S.C. § 112 due to the term "said task definition" lacking a proper antecedent basis in the claims.

Appellants admit that the claim language is not necessarily as clear as it could be. When practicing the invention, multiple task definitions are sent to multiple CPUs, but each task definition has the same properties and is sent only when the data elements it references are available. Despite variously using "tasks," "task definitions" (twice), "a task definition for each task,"

and "all parts of a task definition" prior to claiming "said task definition," Appellants submit that the claim language is sufficiently clear to refer to the immediately preceding form of the term "all parts of a task definition" as used in the term "all parts of a task definition and data elements referenced by said task definition." Even though the claim language may not be as precise as the examiner might desire, the examiner never suggested claim language to Appellants to improve the clarity or precision of the language used, but instead improperly rejected the claims despite the mode of expression selected by Appellants satisfying the statutory requirement, in direct contravention of M.P.E.P. § 2173.02.

Indeed, under M.P.E.P. § 2173.05(e), a claim is indefinite when it contains words or phrases whose meaning is unclear. The examples in the M.P.E.P. suggest that a lack of clarity could arise where a claim refers to "said lever" or "the lever," where the claim contains no earlier recitation or limitation of a lever and where it would be unclear as to what element the limitation was making reference. Similarly, if two different levers are recited earlier in the claim, the recitation of "said lever" in the same or subsequent claim would be unclear where it is uncertain which of the two levers was intended. No such problems exist in the present claims; "a task definition" is claimed prior to "said task definition" and repeated use of "a task definition" in the phrase "sending a task definition for each task from the master CPU to one of a plurality of slave CPUs when all parts of a task definition and data elements referenced by said task definition are available at said master CPU" does not suggest that the

repeated use of "a task definition" refers to different task definitions.

In view of this, Appellants submit that the scope of the claims would be reasonably ascertainable by those skilled in the art, and therefore the claims are not indefinite. *Ex parte Porter*, 25 USPQ2d 1144, 1145 (Bd. Pat. App. & Inter. 1992) ("controlled stream of fluid" provided reasonable antecedent basis for "the controlled fluid").

In view of the above-cited reasons, Appellants submit that claims 1-23 are definite and respectfully request reconsideration and withdrawal of the rejections. Appellants further note that claims 2-3, 5, and 14-17 have not been rejected in view of the prior art and are thus admittedly allowable upon being found definite.

Reply to Examiner's Response to Arguments

In discussing Appellants arguments, the Examiner's Answer mailed July 13, 2005 erroneously alleges that "'said task definition' is equally applicable to be referring to other 'task definition' recitations previously found in the claims" (emphasis added). This argument misrepresents the claim language since there is only one other recitation of "task definition," the one that occurs in the term "a task definition for each task," which, as illustrated below, inherently is related to plural task definitions.

Indeed, Appellants submit that all previous recitations in the claims, except for one, refer to plural task definitions since the claim defines "tasks" as plural, mentions "task definitions" twice as plural, and recites "a task definition for each task," which is inherently plural since there are plural tasks. However, the context of only

one prior occurrence of "task definition" is singular: "all parts of a task definition and data elements referenced by said task definition."

The argument in the Examiner's Answer, that the examiner never suggested claim language to Appellants under M.P.E.P. 2173.05(e) to improve the clarity or precision of the language used because "it is unclear as to which" task definition Appellant intended "said task definition" to refer to, is disingenuous. Appellant has responded repeatedly to this rejection to submit what "said task definition" refers to, yet the examiner still did not suggest any claim language. Furthermore, Appellants submit that a phrase relating "all parts of" and "data elements referenced by" are clearly related to the same task definition.

The Examiner's Answer further argues that *Ex Parte Porter* "does not appear to be relevant to the instant application as Ex parte Porter deals with claim scope issues..." Appellant submits that the Examiner's Answer clearly fails to understand that the ultimate reasoning for a rejection under 35 U.S.C. §112 based upon a "lack of antecedent basis" is, in fact, based on the scope of the claim. Indeed, with respect to antecedent basis, *Ex parte Porter* has held that if the scope of a claim would be reasonably ascertainable by those skilled in the art, then the claim is not indefinite.

Claim Rejections - 35 USC §103

Grounds 2

Claims 1, 4, 6-13, and 18-23 were rejected under 35 U.S.C. 103 are being obvious over the publication to Smith

et al. (1996) in view of the publication to Altschul et al. (1990) and U.S. Patent No. 5,862,325 to Reed et al.

To establish a *prima facie* case of obviousness, three basic criteria must be met (See M.P.E.P. Section 2143). First, there must be some suggestion or motivation, either in the references themselves or in the knowledge generally available to one of ordinary skill in the art, to modify the reference or to combine reference teachings. *In re Fine*, 837 F.2d 1071, 5 USPQ2d 1596 (Fed. Cir. 1988); *In re Jones*, 958 F.2d 347, 21 USPQ2d 1941 (Fed. Cir. 1992).

Second, there must be a reasonable expectation of success. This requirement is primarily concerned with less predictable arts, such as the chemical arts.

Finally, the prior art must teach or suggest each and every limitation of the claimed invention, as the invention must be considered as a whole. *In re Hirao*, 535 F.2d 67, 190 U.S.P.Q. 15 (C.C.P.A. 1976).

The teaching or suggestion to make the claimed combination and the reasonable expectation of success must both be found in the prior art, not in Apellants' disclosure. *In re Vaeck*, 947 F.2d 488, 20 USPQ2d 1438 (Fed. Cir. 1991).

No Motivation to Combine

In the present case, none of these criteria have been met in the Office Action. First, there is no suggestion or motivation, either in the references themselves or in the knowledge generally available to one of ordinary skill in the art, to modify the search launcher interface of Smith et al. or combine it with Altschul et al. and Reed et al.

M.P.E.P. 2141.02 requires that an invention be considered as a whole. The present invention, *as a whole*,

is drawn to a method or system for comparing a query dataset N to a subject dataset M using not only a network, but a *distributed computing platform*. A client computer in the claimed system and method divides the query dataset N into n_N data elements having a size within a specified range, divides the subject dataset M into n_M data elements having a size within said specified range, and determines a number of tasks for an entire comparison of datasets N and M as $n_N \times n_M$. The client computer then sends all data elements and task definitions to a master CPU of a master-slave distributed computing platform, and the master CPU sends a task definition and its associated data elements for each task to one of a plurality of slave CPUs of the distributed computing platform. The slave CPUs of the distributed computing platform perform the tasks (inherently in parallel) and return the results to the master CPU.

In contrast, *none* of Smith et al., Altschul et al. or Reed et al. even mentions distributed computing. In making the rejection, the Office Action erroneously looks to *Merriam-Webster* for the definition of "system" instead of looking to the broadest reasonable interpretation *consistent with the specification* as required by M.P.E.P. 2111. Use of *Merriam-Webster* for the definition of "system" is not consistent with the distributed computing platform disclosed in the specification.

M.P.E.P. 2141.02 further requires that the prior art be considered as a whole, including portions that teach away from the invention. Smith et al., as a whole, *teaches against* the present invention in teaching the use of a batch system that processes various sequence searches

serially "one at a time" at a single site (the BCM Search Launcher server, see Abstract, lines 14-17 and page 461, column 2, discussing batch processing) instead of in parallel at multiple slave CPUs, as found in the present invention. Smith et al. is merely a client-server system for providing a search launcher WWW interface and merely provides access to existing WWW services on remote servers. No matter how the Office Action twists or mischaracterizes Smith et al. (i.e., "Smith et al. describes … promoting a distributed information space by filling out an HTML form…"), it is a fact that neither the client nor the BCM server include any step or software for splitting up a NxM dataset comparison into $n_N \times n_M$ tasks. Likewise, it is a fact that client search requests in Smith et al. are processed serially and that each search request is sent to a single remote site. A fair reading of Smith et al. illustrates that the disclosed system is merely a WWW gateway to pre-existing search services and that it can perform some pre-processing in the form of batch entry and post-processing in the form of adding links to results. It does nothing to solve the problems existing in the prior art, such as (1) that sequence-to-database comparisons (as illustrated in fig. 1 of Smith et al.) require large RAM requirements for efficient processing or (2) that typical BLAST queries over a network involve sending inefficient ASCII (256-bit) characters (as illustrated by the "cut and paste" sequence entry disclosed by Smith et al.).

Likewise, Altschul et al., as a whole, *teaches away* from the present invention by teaching dataset-to-dataset comparison on a single machine (i.e., "a shared memory version of BLAST…loads the compressed DNA file into memory

once" is the only disclosed technical performance enhancement). Although Altschul et al. discloses the comparison of two random sequences *n* and *m*, it nowhere suggests dividing the problem further, let alone dividing it into tasks for different computers to solve, as erroneously asserted by the Office Action.

The Office Action's citation to Reed et al. borders on the ridiculous. Reed et al. has nothing to do with bioinformatics. It has nothing to do with dataset comparisons. Indeed, it has nothing to do with solving large computational problems with distributed computing (but rather deals with information distribution). Reed et al. is drawn to an "automated communications system [that] operates to transfer data, metadata and methods from a provider computer to a consumer computer through a communications network." The disclosed compression in col. 57 is for word processing documents with PKZIP, not the databases of col. 14 as the Office Action implies. Like the other references, it *teaches away* from the present invention since, as the cited paper to Matsumoto et al. teaches on page 44, "if one applies the standard text compression software such as compress or gzip, they cannot compress DNA sequences, but only expand the file with more than two bits per symbol." The present invention applies standard redundancy reduction data compression to an efficiently packed data element to avoid this issue, not the raw ASCII data that Smith et al. and Reed et al. seek to transmit over networks.

The stated motivation for the combination in the Office Action, i.e., that "it would have been obvious one having ordinary skill in the art at the time the invention

was made to compress data (as stated by Altschul et al. and Reed et al.) and looping processes [sic] (as stated by Reed et al.) in order to offer enhanced, integrated, easy-to-use, and time-saving techniques to a large number of useful molecular biology database search and analysis services for organizing and improving access to these tools for Genome researchers worldwide (Smith et al., page 459, col. 1, third paragraph to col. 2, first paragraph)" is not only incomprehensible, but it further is *completely unrelated to limitations of the claimed invention*. It is clearly an improper hindsight reconstruction, not even of the claimed invention, but merely for the purpose of combining the disparate references that the Examiner found that use appropriate words like "BLAST," "server," "network," "distributed," "database," and "compression," which apparently turned up in the required electronic text searches.

Indeed, the Office Action has completely failed at making a *prima facie* case of obviousness under *Graham v. Deere* since it has failed to identify or evaluate *any* of the differences between the claimed invention and the prior art.

No Reasonable Expectation of Success

One of ordinary skill in the art could not reasonably be expected to find Applicant's claimed invention for comparing large datasets obvious in view of a plurality of references that provide no guidance on handling large datasets or processing them in parallel over a network. Indeed, if the compression teaching suggested by the Office Action were implemented (PKZIP compression of ASCII DNA

data), the network would be saturated (and fail) due to the *expanded* file sizes that would result therefrom.

All Claim Limitations Not Shown

Claims 1, 12-13, and 23

Smith et al. teaches the running of sequence-to-database searches, but fails to teach or fairly suggest numerous claim limitations required by all of the claims, including at least the following found in claims 1 and 13:

- dividing said query dataset N into n_N data elements having a size within a specified range [at client CPU];
- dividing said subject dataset M into n_M data elements having a size within said specified range[at client CPU];;
- determining a number of tasks for an entire comparison of datasets N and M as $n_N \times n_M$[at client CPU];;
- sending all data elements and task definitions to a master central processing unit (CPU) of a master-slave distributed computing platform,

  wherein task definitions comprise at least one comparison parameter, at least one executable element capable of performing comparisons, a query data element identification(ID)/descriptor, and a subject data element ID/descriptor, and

  wherein data elements are sent alternately from query and subject data elements;
- sending a task definition for each task from the master CPU to one of a plurality of slave CPUs when all parts of a task definition and data elements referenced by said task definition are available at said master CPU;
- sending data elements referenced by said task definition to said slave CPU; and
- performing each task on a slave CPU.

Appellant respectfully submits that selection of a sequence to "clip and paste" into the HTML input form of Smith et al. is not a *division* of a query dataset N, but rather a specification of dataset N. No datasets in Smith et al. are ever divided, no tasks (*plural* for a single N-M comparison) are determined, and no subject dataset elements are ever sent to a Master CPU. Despite the allegation in the Examiner's Answer that "['tasks are determined'] is not a limitation of the instant claims," both claim 1 and 13 clearly include "determining a number of tasks for an entire comparison of datasets N and M as $n_N \times n_M$."

Altschul et al. fail to disclose any of the limitations missing from Smith et al. It merely discloses the basic BLAST algorithm for sequence comparison, i.e., comparing one sequence with another sequence, or for searching a database. Like Smith et al., Altschul et al. at least fail to disclose or suggest dividing sequence comparison problems into discrete segments for processing on a plurality of CPUs, let alone any specific method of doing this task.

Reed et al. also fail to remedy any of the defects of the Smith et al. and Altschul et al. references and is completely unrelated to the present invention.

As a whole, none of the cited prior art teaches or fairly suggests dividing the problem of comparing datasets M and N into $n_N \times n_M$ comparisons of data elements from N with data elements from M as presently claimed. For at least these reasons, Appellants submit that the claims are allowable over the prior art and request reconsideration and allowance of the claims.

Claim 4

Despite teaching the use of redundancy reduction data compression, the prior art fails to disclose or teach the combined steps of:

- stripping all metadata from data;
- packing said data into an efficient structure;
- creating an index for said data and packing said index and said data in an uncompressed data structure; and
- compressing said uncompressed data structure into a data element using a redundancy reduction data compression method.

While the Examiner's Answer does not reject the substantially similar claim 16, it is further submitted that claim 4 is allowable for the same reasons as claim 16, and *vice versa*.

Claims 6 and 18

As submitted above, the "looping" of Reed et al. is unrelated to the present invention. As such, Appellant submits that the prior art fails to disclose or teach the combined steps of:

- uncompressing and unpacking data from said query and subject data elements;
- looping through query sequences from said query data element to perform setup, preprocessing and table generation for each row of comparisons;
- looping through subject sequences from said subject data element and, for each pair of query and subject sequences, performing a comparison using said executable element and finding results based on said at least one comparison parameter; and
- storing minimal information that will allow

reconstruction of said result.

Claims 7 and 19

Despite allegations that a window in fig. 2 of Smith et al. meets the limitations of these claims, Smith et al. fails to disclose or suggest the combination of:

- storing index information for said query and said subject sequence;
- storing bounds information for start and stop of said query and subject sub sequences;
- storing data that quantify fulfillment of significance criteria for a significant match; and
- storing an efficiently encoded representation of alignment between said bounds corresponding to a high-scoring segment pair.

Claim 8

Despite allegations that a window in fig. 2 of Smith et al. meets the limitations of this claim, Smith et al. fails to disclose or suggest:

- storing a seed point and sum-set membership for each alignment for Basic Local Alignment Search Tool (BLAST).

Claims 9-10 and 20-21

Despite allegations about what Smith et al. discloses, Smith et al. clearly does not disclose determining tasks and therefore cannot possibly store results for tasks, including the claimed:

- storing task results in a task result file, said file including query and subject sequence data and metadata corresponding to the task that the results came from,

metadata for the subject sequence, the partial subject sequence data corresponding to the subject bounds of the significant alignment result, and any other results data for each result in the task results

Claims 11 and 22

Despite allegations about what Smith et al. discloses, Smith et al. does not disclose splitting dataset comparisons into smaller tasks and therefore has no need to process multiple results, including the claimed:

- concatenating results from all BLAST reports.

Reply to the Examiner's Response to Arguments
Final Office Action

With respect to the motivation or suggestion to combine references, the Final Rejection states that "Smith et al. state the problem of hindering efficient use as well as improving and simplifying access and sources which is a proper motivation to combine." Regardless, Smith et al. only suggests an improved interface with batch processing that *teaches against* the distributed processing of the present invention.

With regard to the claims not reciting the phrase "distributed computing," Appellants note that the claimed invention and the prior art must be considered *as a whole*. As a whole, the present claims define a distributed computing platform and method, whereas the cited prior art does not.

Likewise, the interpretation of the claimed "master-slave distributed computing platform" to be covered by the "system" of Smith et al. is clearly inconsistent with the specification, in contravention of M.P.E.P. 2111.

The Examiner's arguments related to "parallel use" not being recited in the claims again fails to address the claims as a whole, which claim dividing a comparison into tasks, sending the tasks to be computed on a plurality of CPUs, and returning task results - the very definition of parallel computing.

In regard to the Examiner's argument that not all limitations need to be found in each reference, Appellants note that (1) Smith et al. fails to teach or suggest that any dataset-to-dataset comparison is performed on more than one machine (the portal merely provides access to existing services), (2) Altschul et al. teaches dataset-to-dataset comparison on a single machine, and (3) Reed et al. has nothing to do with dataset comparisons.

With regard to the teachings of Reed et al., the Examiner misses the point that there is no reason to combine Reed et al. with the bioinformatics references absent impermissible hindsight. Smith et al. teaches basic clip-n-paste submissions of queries to run against external databases. Ordinary results are returned. There is no suggestion of compression and there is no metadata, so there is no reason to strip it. Likewise, Altschul et al. teaches a basic sequence-to-sequence comparison on a local machine. Again, no metadata or compression is needed or desirable. The only reason to look to the diverse art of Reed et al. is the Appellants' disclosure.

With regard to Smith et al. on pages 15-16, the quoted portions of the Office Action merely repeat the unsupported contentions in a long narrative that fails to match claim limitations with specific portions of the prior art. For example, if the Examiner were to try a proper comparison

between the claims and the prior art under Graham v. Deere, he would immediately see that a client CPU in Smith et al., consisting of a computer with a browser, fails to include instructions for dividing datasets N and M, as required.

No Motivation to Combine

In responding to Appellant's arguments related to a lack of a motivation to combine references, the Examiner's Answer cites to alleged motivation in paragraphs 1 & 2 of page 455 of Smith et al. related to the desirability of an "improved interface to molecular biology-related search and analysis services on the WWW." As Smith et al. is the primary reference in the obviousness rejection, it is unclear *why* it would be desirable to modify Smith et al. with any secondary references when Smith et al. *already includes* the desirable features related to improvements to analysis server sites, simplified access, and improved analysis resources.

Again, as previously submitted, M.P.E.P. 2141.02 requires that an invention be considered *as a whole*. The present invention, *as a whole*, is drawn to a method or system for comparing a query dataset N to a subject dataset M using not only a network, but a *distributed computing platform*. A client computer in the claimed system and method divides the query dataset N into n_N data elements having a size within a specified range, divides the subject dataset M into n_M data elements having a size within said specified range, and determines a number of tasks for an entire comparison of datasets N and M as $n_N \times n_M$. The client computer then sends all data elements and task definitions to a master CPU of a master-slave distributed computing platform, and the master CPU sends a task definition and

its associated data elements for each task to one of a plurality of slave CPUs of the distributed computing platform. The slave CPUs of the distributed computing platform perform the tasks (inherently in parallel) and return the results to the master CPU.

In contrast, *none* of Smith et al., Altschul et al. or Reed et al. even mentions distributed computing. In making the rejection, the Final Office Action erroneously looks to *Merriam-Webster* for the definition of "system" instead of looking to the broadest reasonable interpretation *consistent with the specification* as required by M.P.E.P. 2111. Looking to *Merriam-Webster* for the definition of "system" is not consistent with the distributed computing platform disclosed in the specification.

In addressing this issue, the Examiner's Answer points out that "the words of a claim must be given their plain meaning unless applicant has provided a clear definition in the specification" and that "no clear and concise definition of 'distributed computing' was set forth in the specification." Appellant notes, however, that these arguments only serve to *amplify* the completely fallacious nature of the rejection.

First, the Examiner's Answer did not interpret "distributed computing" or "distributed computing platform," but rather "system." Second, claim 1 does not even mention "system," the term allegedly given its plain meaning from *Merriam-Webster*, but rather "a master central processing unit (CPU) of a master-slave distributed computing platform" and a "plurality of slave CPUs." Claim 13 uses "system" in the preamble, but still claims "a master central processing unit (CPU) of a master-slave

distributed computing platform" and a "plurality of slave CPUs." Indeed, the explicit language of the rejection illustrates that the term "system" was construed relative to Smith et al., not the claim terms: "Smith et al. discloses...which represents a 'system' which is defined by the online Merriam-Webster dictionary as..." (see Examiner's Answer, p. 5, l. 9-10).

Instead of interpreting the language of the claims, the Examiner's Answer confirms that it is Smith et al. that has been interpreted. The rejection clearly interprets a term from Smith et al., not a *claim term*, and is thus improper. Indeed, Appellant submits that the plain meaning of "a master central processing unit (CPU) of a master-slave distributed computing platform" and a "plurality of slave CPUs" is as previously submitted on March 9, 2004: the terms "master CPU," "slave CPU," and "master-slave distributed computing platform" have established meanings to one of skill in the art; A master CPU controls and directs the actions (tasks) of the slave CPUs.

In regard to Appellant's argument that Smith et al., as a whole, *teaches against* the present invention in teaching the use of a batch system that processes various sequence searches serially "one at a time" at a single site (the BCM Search Launcher server, see Abstract, lines 14-17 and page 461, column 2, discussing batch processing) instead of in parallel at multiple slave CPUs, as found in the present invention, the Examiner's Answer alleges that "the instant claims do not recite parallel use limitations." This allegation is erroneous.

The claims recite "sending a task definition for each task from the master CPU to one of a plurality of slave

CPUs," "sending data elements referenced by said task definition to said slave CPU," "performing each task on a slave CPU," and "returning task results for each task to said master CPU" to inherently define parallel processing. Indeed, the purpose of a distributed computing platform is to distribute the problem being solved to a plurality of computers to have multiple parts of the problem solved in parallel instead of sequentially.

The Examiner's Answer fails to address the argument that Smith et al. is merely a client-server system for providing a search launcher WWW interface and merely provides access to existing WWW services on remote servers. No matter how the Examiner's Answer twists or mischaracterizes Smith et al. (i.e., "Smith et al. describes... promoting a distributed information space by filling out an HTML form..."), it is a fact that neither the client nor the BCM server include any step or software for splitting up a NxM dataset comparison into $n_N \times n_M$ tasks for computation on a plurality of slave CPUs. Likewise, it is a fact that client search requests in Smith et al. are processed *serially* and that each search request is sent to a single remote site. A fair reading of Smith et al. illustrates that the disclosed system is merely a WWW gateway to pre-existing search services and that it can perform some pre-processing in the form of batch entry and post-processing in the form of adding links to results. It does nothing to solve the problems existing in the prior art, such as (1) that sequence-to-database comparisons (as illustrated in fig. 1 of Smith et al.) require large RAM requirements for efficient processing or (2) that typical BLAST queries over a network involve sending inefficient

ASCII (256-bit) characters (as illustrated by the "cut and paste" sequence entry disclosed by Smith et al.).

With respect to Appellant's argument that the secondary reference Altschul et al., as a whole, *teaches away* from the present invention by teaching dataset-to-dataset comparison on a single machine, the Examiner's Answer apparently argues that the primary reference Smith et al. teaches "the ability to use multiple computers." While it is unclear why the ability to use multiple computers such as by having "a batch client interface for Unix and Macintosh computers" (abstract) or "using a plurality of servers that can return search results" (allegedly at p. 455) as cited in the Examiner's Answer teaches anything except batch or serial processing, it does not overcome the fact that, as a whole, Altschul et al. *teaches away* from the presently-claimed distributed computating.

The Examiner's Answer further erroneously alleges that Altschul et al. teaches "division of a dataset into n_N x n_M elements." As previously submitted, although Altschul et al. discloses the comparison of two random sequences *n* and *m*, it nowhere suggests dividing the problem further, let alone dividing it into tasks *for different computers to solve*. The Examiner's Answer is incorrect in alleging that:

> "Altschul et al. describe performing BLAST with two random sequences (data sets N and M) of lengths *m* and *n* in order to determine the probability of finding a segment pair with a score greater than S (cut off score, page 404, col. 2, second paragraph) where $y=Kmn\ e^{-\lambda S}$ (page 405, col. 2, equation 1 and fourth paragraph) which represents determination of a number of tasks or points in a matrix in a comparison of two datasets (i.e., N and M) involving a

> multiplication of data elements n_N x n_M, as stated in the instant claims 1 and 13."

The cited equation relates to the probability of finding a segment pair with a score greater than S where $S=1-e^{-y}$ and has nothing to do "a multiplication of data elements n_N x n_M." In fact, Altschul et al. merely discloses the use of BLAST for comparisons "in a variety of contexts including straight-forward DNA and protein sequence database searches, motif searches, gene identification searches and in the analysis of multiple regions of similarity in long DNA sequences" (see abstract). Database searches are dataset (search term) to dataset (target database). Likewise, sequence (first DNA sequence) to sequence (second DNA sequence) are dataset to dataset comparisons. Altschul et al. does not teach or suggest *any* subdivision of these methods for handling larger datasets.

With respect to the citation of the non-analogous art of Reed et al., Appellant has no idea how the "looping processes" discussed in the Examiner's Answer relate to the claim limitations. The claim 6 and 18 require that each slave CPU perform "looping through query sequences from said query data element to perform setup, preprocessing and table generation for each row of comparisons," and "looping through subject sequences from said subject data element and, for each pair of query and subject sequences, performing a comparison using said executable element and finding results based on said at least one comparison parameter."

In contrast to this, Reed et al. has nothing to do query sequences or the setup, preprocessing, or table generation used in sequence comparison. The "looping" of

Reed et al. is related to the publication of an entire object, by a query to determine object flag status (to find out what needs to be published) a first looping though the communication objects to read the associated recipients and a second loop to generate a communication object instance for each recipient.

While the Examiner's Answer seems to suggest that Reed et al. is analogous since it deals with computer data, Appellant notes that Reed et al. has nothing to do with computerized analysis, computation, or master/slave computation, but rather deals with data communication and processing. Even within class 709, the Patent Office recognizes that distributed data processing (709/201 where Reed et al. is classified) is different from master/slave computer controlling (709/208). Reed et al. is concerned with communicating data between a provider computer and a consumer computer. It has nothing to do with solving the problems encountered by those of ordinary skill in the art of bioinformatics when faced with the problems of large dataset comparisons, and therefore, one of ordinary skill in the art of bioinformatics would not look to Reed et al. for any teachings with respect to solving bioinformatics problems.

A prior art reference is analogous if the reference is in the field of Appellant's endeavor or, if not, the reference is reasonably pertinent to the particular problem with which the inventor was concerned. *In re Oetiker*, 977 F.2d 1443, 1446, 24 USPQ2d 1443, 1445 (Fed. Cir. 1992). In the present case, Reed et al. is not from the field of bioinformatics or dataset comparison and is not pertinent

to the problems of solving large dataset comparisons or the problems of a distributed computing platform.

While the Examiner's Answer is correct in stating that the reasons to combine references need not be the same as Appellants, the reasons to combine must come from the prior art and not from hindsight. "There are three possible sources for a motivation to combine references: the nature of the problem to be solved, the teachings of the prior art, and the knowledge of persons of ordinary skill in the art." *In re Rouffet,* 149 F.3d 1350, 1357, 47 USPQ2d 1453, 1457-58 (Fed. Cir. 1998)

The stated motivation for the combination in the Examiner's Answer is that "it would have been obvious one having ordinary skill in the art at the time the invention was made to compress data (as stated by Altschul et al. and Reed et al.) and to use looping processes (as stated by Reed et al.) in order to offer enhanced, integrated, easy-to-use, and time-saving techniques to a large number of useful molecular biology database search and analysis services for organizing and improving access to these tools for Genome researchers worldwide (Smith et al., page 459, col 1, third paragraph to col. 2, first paragraph)." However, none of this prior art teaches or fairly suggests use of a master/slave distributed computing platform and the stated rejection lacks any reason for modifying the prior art to include this limitation.

Arguably, because Smith et al. teaches access to NCBI BLAST, it already includes the BLAST teachings of Altschul et al. and is completely clear that a combination of the two references/teachings is access to the WWW NCBI BLAST tool through the gateway of Smith et al. Since Smith et al.

teaches the gateway for customer access to bioinformatics tools through the WWW, it is unclear why one would look to the data distribution and access methods of Reed et al.

No Reasonable Expectation of Success

The Examiner's Answer dismissed Appellant's arguments that one of ordinary skill in the art could not reasonably be expected to find Applicant's claimed invention for comparing large datasets obvious in view of a plurality of references that provide no guidance on handling large datasets or processing them in parallel over a network, by stating that "the instant claims do not recite large datasets or parallel processing."

In response, Appellant submits that the problem solved (large dataset comparisons) and inherent properties (datasets split into *smaller* tasks for computation are inherently *large* and parallel processing is inherent in the claimed distributed computing platform) are part of the invention "as a whole" inquiry under 35 U.S.C. §103 (see M.P.E.P. §2141.02).

Response to Examiner's Answer of July 13, 2005

The Examiner's Answer alleges that "Smith et al. teach a distributed computing platform." This is incorrect. Smith et al. teaches a search launcher for molecular-biology related search and analysis services available on the WWW (Internet). The alleged broad interpretation of "distributed computing platform" in the Examiner's Answer is not a reasonable interpretation consistent with the specification of the application as required by M.P.E.P. §2111.

With regard to the limitations in the specification not being read into the claims (page 15, lines 4-5 of the

Examiner's Answer), Appellant notes that the claims already recite a "distributed computing platform" and that the broadest *reasonable* interpretation of the claim terms must be *consistent with the specification* under M.P.E.P. 2111. In the present case, the interpretation of the claimed "master-slave distributed computing platform" as somehow being covered by the "system" of Smith et al. is clearly inconsistent with the specification, in contravention of M.P.E.P. 2111.

If the Examiner's Answer had followed the guidelines of M.P.E.P. 2111.01(II) quoted on page 15 of the Examiner's Answer, it would have produced a standard definition for "distributed computing platform," not "system." Furthermore, the definition of system used by the examiner fails to consider the difference between "distributed" and "distributing."

The definition applied in the rejection is: "a group of devices or artificial objects or an organization forming a network especially for distributing something or serving a common purpose <a telephone *system*><a heating *system*><a highway *system*><a data processing *system*>." This definition fails to suggest any particular *arrangement* of the system, whereas a "distributed computing platform" requires that the computing is done in a *distributed*, i.e., *spread-out, manner*.

The Examiner's Answer repeatedly dismisses Appellant's arguments as "conclusory allegation[s] and unpersuasive without any sound reasoning," yet it is the Examiner's Answer that relies upon conclusory allegations and unsound reasoning. The Examiner's Answer alleges "Smith et al. teach a distributed computing platform," but this is

conclusory. Any fair reading of Smith et al. illustrates that it has nothing to do with distributed computing platforms and instead discloses a collection of Web pages for accessing bioinformatics services available on the WWW.

Likewise, the Examiner's Answer relies upon illogical reasoning such as use of a dictionary definition of a term *in the prior art* - "system" in Smith et al. - to somehow interpret the meaning of the claim term "distributed computing platform."

Furthermore, Appellant's arguments have *not* been conclusory and without sound reasoning. Appellant has repeatedly submitted that: Smith et al. is merely a client-server system for providing a search launcher WWW interface and merely provides access to existing WWW services on remote servers; neither the client nor the BCM server include any step or software for splitting up a NxM dataset comparison into $n_N \times n_M$ tasks for computation on a plurality of slave CPUs; client search requests in Smith et al. are processed *serially* and that each search request is sent to a single remote site; and a fair reading of Smith et al. illustrates that the disclosed system is merely a WWW gateway to pre-existing search services and that it can perform some pre-processing in the form of batch entry and post-processing in the form of adding links to results.

With respect to the motivation or suggestion to combine references, the Examiner states that "Smith et al. state the problem of hindering efficient use as well as improving and simplifying access and sources which is a proper motivation to combine." Regardless of whether this provides any motivation to look to Altschul et al. or Reed et al. (it doesn't since it only states a beneficial

quality of the primary reference), Smith et al. only suggests an improved interface with *serial batch processing* that *teaches against* the *distributed processing* of the present invention.

With regard to the claims not reciting the phrase "distributed computing" (page 14, lines 5-6 of the Examiner's Answer), Appellant notes that the remaining portion of the sentence in the Examiner's Answer, stating what the claim recites, *clearly includes the phrase.* Indeed, all the claims require a "distributed computing platform" with a "master CPU" and a plurality of "slave CPUs."

The Examiner's Answer, spanning pages 15-16 and again at page 19, lines 8-9, dismisses arguments related to "parallel use" since the term is not recited in the claims. Again, this fails to address the claims *as a whole*. The claims require dividing a dataset-dataset comparison into a plurality tasks, sending the tasks to be computed on a plurality of CPUs, and returning task results - the very definition of parallel computing.

Page 19, lines 12-15 of the Examiner's Answer states that "Smith et al. describes...a network for distributing [and] using the Internet for...promoting a distributed information space." Regardless of the veracity of this statement, "a network for distributing" and a "distributed information space" are clearly not a "distributed computing platform" consistent with the specification.

In regard to the Examiner's argument that not all limitations need to be found in each reference, Appellants note that (1) Smith et al. fails to teach or suggest that any particular dataset-to-dataset comparison is performed

on more than one machine (the portal merely provides access to existing services, and providing the same or other dataset-dataset comparisons on other machines, i.e., the "dataset comparisons at multiple sites," is not the same), (2) Altschul et al. teaches dataset-to-dataset comparison on a single machine, and (3) Reed et al. has nothing to do with dataset comparisons.

With regard to the teachings of Reed et al., the Examiner misses the point that there is no reason to combine Reed et al. with the bioinformatics references absent impermissible hindsight. Smith et al. teaches basic clip-n-paste submissions of queries to run against external databases. Ordinary results are returned. There is no suggestion of compression and there is no metadata, so there is no reason to strip it. Likewise, Altschul et al. teaches a basic sequence-to-sequence comparison on a local machine. Again, no metadata or compression is needed or desirable. The only reason to look to the diverse art of Reed et al. is Appellants' disclosure.

Indeed, it is arguable that since a primary purpose of Smith et al. is to provide easy access to bioinformatics tools through a standard web browser, and that clip and paste entry is part of what makes it easy to use, that the proposed modification of Smith et al. with Reed et al. to add compression, etc. would destroy the principal of operation (standard browser) and thus be improper under M.P.E.P. 2143.01 since: (i) If proposed modification would render the prior art invention being modified unsatisfactory for its intended purpose, then there is no suggestion or motivation to make the proposed modification, *In re Gordon*, 733 F.2d 900, 221 USPQ 1125 (Fed. Cir. 1984);

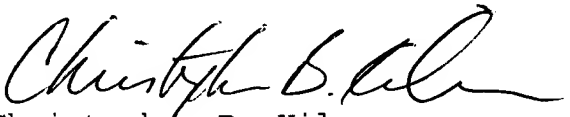
and (ii) If the proposed modification or combination of the prior art would change the principle of operation of the prior art invention being modified, then the teachings of the references are not sufficient to render the claims *prima facie* obvious. *In re Ratti*, 270 F.2d 810, 123 USPQ 349 (CCPA 1959).

Despite finding Appellant's arguments unpersuasive at page 20, lines 4-9, it is undisputable that the portion of the Office Action discussed in Appellant's arguments fails to match *claim limitations* with specific portions of the prior art. The Examiner's Answer fails to address the Appellant's argument that a client CPU in Smith et al., consisting of a computer with a browser, clearly fails to include instructions for dividing datasets N and M, as required.

CONCLUSION

For the above reasons, Appellants respectfully submit that the present claims meet the requirements of 35 U.S.C. 112 and that the Examiner has failed to make out a *prima facie* case of obviousness under 35 U.S.C. 103 with regard to claims 1, 4, 6-13, and 18-23 and asks that the obviousness rejection be reversed.

Respectfully submitted,



Christopher B. Kilner
Registration No. 45,381
Roberts Abokhair & Mardula, LLC
11800 Sunrise Valley Drive,
Suite 1000
Reston, VA 20191
(703) 391-2900

VIII - Claims Appendix

1. A method of comparing a query dataset N with a subject dataset M, comprising:

dividing said query dataset N into n_N data elements having a size within a specified range;

dividing said subject dataset M into n_M data elements having a size within said specified range;

determining a number of tasks for an entire comparison of datasets N and M as $n_N \times n_M$;

sending all data elements and task definitions to a master central processing unit (CPU) of a master-slave distributed computing platform,

wherein task definitions comprise at least one comparison parameter, at least one executable element capable of performing comparisons, a query data element identification(ID)/descriptor, and a subject data element ID/descriptor, and

wherein data elements are sent alternately from query and subject data elements;

sending a task definition for each task from the master CPU to one of a plurality of slave CPUs when all parts of a task definition and data elements referenced by said task definition are available at said master CPU;

sending data elements referenced by said task definition to said slave CPU;

performing each task on a slave CPU; and

returning task results for each task to said master CPU.

2. The method of claim 1, further comprising randomizing sequence order of each dataset if either dataset contains related sequences in a contiguous arrangement.

3. The method of claim 1, further comprising formatting said datasets so as to use exactly the same ambiguity substitutions.

4. The method of claim 1 wherein dividing said datasets into data elements further comprises:

stripping all metadata from data;

packing said data into an efficient structure;

creating an index for said data and packing said index and said data in an uncompressed data structure; and

compressing said uncompressed data structure into a data element using a redundancy reduction data compression method.

5. The method of claim 1, further comprising sending remaining data elements from a more numerous of said datasets to said master CPU followed by all task definitions for otherwise complete tasks if there are fewer data elements from one dataset.

6. The method of claim 1 wherein performing a task on said slave CPU further comprises:

uncompressing and unpacking data from said query and subject data elements;

looping through query sequences from said query data element to perform setup, preprocessing and table generation for each row of comparisons;

looping through subject sequences from said subject data element and, for each pair of query and subject sequences, performing a comparison using said executable element and finding results based on said at least one comparison parameter; and

storing minimal information that will allow reconstruction of said result.

7. The method of claim 6 wherein storing said minimal information comprises:

storing index information for said query and said subject sequence;

storing bounds information for start and stop of said query and subject sub sequences;

storing data that quantify fulfillment of significance criteria for a significant match; and

storing an efficiently encoded representation of alignment between said bounds corresponding to a high-scoring segment pair.

8. The method of claim 7, further comprising storing a seed point and sum-set membership for each alignment for Basic Local Alignment Search Tool (BLAST).

9. The method of claim 7, further comprising storing task results in a task result file, said file including query and subject sequence data and metadata corresponding to the task that the results came from, metadata for the subject sequence, the partial subject sequence data corresponding to the subject bounds of the significant alignment result, and any other results data for each result in the task results.

10. The method of claim 9, further comprising generating a BLAST report for each query data element.

11. The method of claim 10, further comprising concatenating results from all BLAST reports to produce a text file identical to a blastall run of said query and subject datasets.

12. The method of claim 1 wherein said datasets are selected from the group consisting of genomic and proteomic databases.

13. A system for comparing a query dataset N with a subject dataset M, comprising:

a master central processing unit (CPU) of a master-slave distributed computing platform;

a plurality of slave CPUs capable of communication with said master CPU; and

a client CPU with instructions for:

dividing said query dataset N into n_N data elements having a size within a specified range;

dividing said subject dataset M into n_M data elements having a size within said specified range;

determining a number of tasks for an entire comparison of datasets N and M as $n_{N \times} n_M$;

sending all data elements and task definitions to said master CPU of a master-slave distributed computing platform,

wherein task definitions comprise at least one comparison parameter, at least one executable element capable of performing comparisons, a query data element identification (ID)/descriptor, and a subject data element ID/descriptor, and

wherein data elements are sent alternately from query and subject data elements;

said master CPU comprising instructions for:

sending a task definition for each task to one of said plurality of slave CPUs when all parts of a task definition and data elements referenced by said task definition are available at said master CPU; and

sending data elements referenced by said task definition to said slave

CPU; and

said slave CPUs including instructions for:

performing each task; and

returning task results for each task to said master CPU.

14. The system of claim 13, further comprising means for randomizing sequence order of each dataset if either dataset contains related sequences in a contiguous arrangement.

15. The system of claim 13, further comprising means for formatting said datasets so as to use exactly the same ambiguity substitutions.

16. The system of claim 13, wherein said instructions for dividing said datasets into data elements further comprises instructions for:

stripping all metadata from data;

packing said data into an efficient structure;

creating an index for said data and packing said index and said data in an uncompressed data structure; and

compressing said uncompressed data structure into a data element using a redundancy reduction data compression method.

17. The system of claim 13, further comprising instructions for sending remaining data elements from a more numerous of said datasets to said master CPU followed by all task definitions for otherwise complete tasks if there are fewer data elements from one dataset.

18. The system of claim 13, wherein instructions for performing a task on said slave CPU further comprises instructions for:

uncompressing and unpacking data from said query and subject data elements;

looping through query sequences from said query data element to perform setup, preprocessing and table generation for each row of comparisons;

looping through subject sequences from said subject data element and, for each pair of query and subject sequences, performing a comparison using said executable element and finding results based on said at least one comparison parameter; and

storing minimal information that will allow reconstruction of said result.

19. The system of claim 18, wherein said instructions for storing said minimal information comprises instructions for:

storing index information for said query and said subject sequence;

storing bounds information for start and stop of said query and subject sub sequences;

storing data that quantify fulfillment of significance criteria for a significant match; and

storing an efficiently encoded representation of alignment between said bounds corresponding to a high-scoring segment pair.

20. The system of claim 19, further comprising instructions for storing task results in a task result file, said file including query and subject sequence data and metadata corresponding to the task that the results came from, metadata for the subject sequence, the partial subject sequence data corresponding to the subject bounds of the significant alignment result, and any other results data for each result in the task results.

21. The system of claim 20, further comprising instructions for generating a Basic Local Alignment Search Tool (BLAST) report for each query data element.

22. The system of claim 21, further comprising means for concatenating results from all BLAST reports to produce a text file identical to a blastall run of said query and subject datasets.

23. The system of claim 13, wherein said datasets are selected from the group consisting of genomic and proteomic databases.

IX. EVIDENCE APPENDIX

There is no evidence entered by the Examiner that is relied upon by Appellants.

(X) RELATED PROCEEDINGS APPENDIX

There are no related proceedings.